DOCUMENT RESUME

ED 212 263 IR 009 969

AUTHOR	Mischo, William H.
TITLE	A Subject Retrieval Function for the Online Union Catalog. Technical Report.
INSTITUTION	OCLC Online Computer Library Center, Inc., Dublin, Ohio.
REPORT NO	OCLC/DD/TR-81/4
PUB DATE	20 Nov 81
NOTE	31p.
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Computer Oriented Programs; *Information Retrieval; *Online Systems; *Permuted Indexes; *Search Strategies; *Subject Index Terms
IDENTIFIERS	Boolean Logic; Library of Congress; MARC Data Base; OCLC

ABSTRACT

This technical report describes a proposed subject retrieval function for the OCLC Online Union Catalog and examines the inadequacies of Library of Congress Subject Headings (LCSH) for the provision of subject access. The function described uses a non-Boolean search approach over data elements present in the MARC record format. Centered around the creation of additional subject index keys and cross-references by LCSH phrase manipulation and title phrase extraction, it would employ familiar OCLC derived key conventions and simulate a Boolean search capability by building precoodinated term combinations into derived index keys. The proposed approach would offer greatly improved subject access over the traditional card catalog by increasing the average number of subject access points per title from 1.4 to more than 22. Four illustrations and 12 references are included. (Author/RAA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED212263

Report Number: OCLC/DD/TR-81/4
Date: 1981 November 20

Technical Report

on

A SUBJECT RETRIEVAL FUNCTION FOR THE ONLINE UNION CATALOG

by

William H. Mischo

OCLC
Development Division
Library Systems Analysis and Design Department
6565 Frantz Road
Dublin OH 43017

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Lois Yoakam

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

IR009969

Report Number: OCLC/DD/TR-81/4
Date: 1981 November 20

ABSTRACT

This technical report describes a proposed subject retrieval function for the OCLC Online union Catalog. The author examines the inadequacies of Library of Congress Subject Headings (LCSH) in providing subject access to library materials.

The subject retrieval function described in the report uses a non-Boolean search approach over data elements present in the MARC record format.

The subject retrieval function is centered around the creation of additional subject index keys and cross-references from assigned LCSH and title keywords and phrases.

The retrieval function would employ familiar OCLC derived search key conventions and is designed to simulate a Boolean search capability by building precoordinated term combinations into the derived index keys. The proposed approach would offer greatly improved subject access over the traditional card catalog approaches by increasing the average number of subject access points per title from 1.4 to over 22.

Report Number: OCLC/DD/TR-81/4
Date: 1981 November 20

CONTENTS

ABSTRACT iii

LIST OF FIGURES vi

I. INTRODUCTION AND OVERVIEW 1

II. PROBLEMS OF SUBJECT ACCESS 3
A. Specificity 3
B. Currentness 3
C. Phrase Construction 3
D. Indexing Exhaustivity 5
E. Syndetic Structure 5
F. Lack of Guiding Principles in LCSH 6

III. ENHANCING SUBJECT ACCESS 7
A. Rotation of Subject Headings 7
B. Number of Subject Heading Access Points 10
C. Title Word Access 11
D. Overview of Subject Access Enhancements 12
E. Other Measures to Improve Subject Access 13

IV. SYSTEM DESIGN FEATURES 14
A. File Design 14
B. Subject Search Transactions 16
C. Limiting of Subject Search Results 19
D. Charging for Offline Products and Basic Service 20
E. Public Access Enhancements 20
F. Manipulation of Search Keys 21
G. Comparison with Boolean Searching 21
H. Summary of Subject Retrieval Function 22

V. CONCLUSION 24

VI. REFERENCES 25

LIST OF FIGURES

No.	Title	Page
1	Subject Authority Record Format	10
2	Subject Authority File	15
3	Subject Access Data Base	16
4	Subject Retrieval Function	17

Report Number: OCLC/DD/TR-81/4
Date: 1981 November 20

I. INTRODUCTION AND OVERVIEW

Subject searching, because of the vagaries and treacheries of natural language, is an inexact science under all retrieval conditions and systems.[1] The well-documented tendency of library catalog users to give up a search after checking at a first entry point only exacerbates the problems of subject searching schemes.[2,3] Conflicting views exist among librarians concerning the purpose and function of the subject catalog. [4] To some the subject catalog should be a comprehensive, in-depth index to library materials, providing access to chapters in books and even periodical articles. To others the subject catalog is merely a convenient instrument for locating some relevant materials on a particular subject, and must be complemented by subject bibliographies and indexes to obtain complete subject control of library materials. The Library of Congress Subject Headings (LCSH) scheme represents this second approach, providing brief subject control over library materials. However, because of shortcomings and historical problems, LCSH does not adequately supply even this moderate subject control. The inadequacies of subject heading access via the de facto national subject scheme--Library of Congress Subject Headings (LCSH)--and the resulting limitations in subject access to library monographic materials is considered the most serious deficiency of present bibliographic control.[5,6,7,8]

This technical report will, first, examine the inadequacies of subject access via LCSH, and second, describe a subject retrieval function for the OCLC Online Union Catalog (OLUC), using only elements in the MARC record. This subject retrieval function utilizes a non-Boolean approach and is centered around the creation of additional subject index keys by LCSH phrase manipulation and title phrase extraction.

The subject retrieval function described in this technical report provides greatly increased subject access over the traditional card catalog approaches. The proposed system would provide approximately ten times the number of subject access points present in the card catalog. The proposed subject retrieval function uses familiar OCLC searching conventions and is tailored to search key inquiries of the form 5, O-5, O-2, or anana, [[anana], [an]] (both commas being required elements of the search key).

The proposed subject retrieval system features:

(1) Access by subject heading keywords and subfield 'a' term pairs;

(2) Added precoordinated subject headings generated by rearranging the order of heading subfields so that all subfields are given lead term status;

(3) Title keyword access points which refer the user to the subject heading(s) assigned to works in which the particular word/phrase appears in the title or subtitle;

(4) The cross-reference structure ("see" and "see also" references) present in LCSH; and

(5) The capability of qualifying search results by holding library, year and language of publication, bibliographic level and type, and other data from the 008 fixed field leader elements.

In total, an estimated 22 subject access points--all of which can be retrieved by initial keyword--will be established per title.

II. PROBLEMS OF SUBJECT ACCESS

The weaknesses of the LCSH scheme are manifold. Briefly listed, these problems include: (1) lack of specificity in subject headings, (2) lack of currentness in subject headings, (3) inconsistent and awkward phrase construction, (4) lack of indexing exhaustivity, (5) lack of systematic cross-referencing or syndetic structure, and (6) an overall lack of guiding principles.

A. Specificity

One shortcoming of LCSH is its lack of specificity in assignment of subject headings. For example, titles on "Lebesque integration," "fuzzy sets," and "semisets" are all assigned the general heading SET THEORY. A book on California's Proposition 13 can only be found under REAL PROPERTY TAX--CALIFORNIA. Works with the phrase "military-industrial complex" in the title are found under the headings UNITED STATES--MILITARY--DEFENSE or INDUSTRY AND STATE--UNITED STATES. Titles on the "space shuttle" are assigned the subject heading REUSABLE SPACE VEHICLES. In all the above cases, no cross-references were initially established from the more specific terms to the subject headings representing them. As a matter of policy, LC does not include upward, or narrower-to-broader cross-references although this rule has been relaxed somewhat in recent years.[9]

In large catalogs, the assignment of general headings, such as EDUCATION, results in headings with large numbers of postings. Using the main entry to secondarily arrange titles with the same subject headings further prohibits efficient subject retrieval for library users interested in locating the most recent works on a subject.

B. Currentness

The presence of archaic and discriminatory terminology in LCSH is a well-documented problem. EUROPEAN WAR--1914-1918 is used rather then World War I, BIOLOGICAL CHEMISTRY for biochemistry, MISCEGENATION in place of interracial dating. Other terms such as YELLOW PERIL and MAMMIES serve as further examples of the problems in outdated LCSH terminology.

C. Phrase Construction

Closely related to the specificity problem is the presence in LCSH of inconsistent and awkward phrase constructionsConsider the following subject headings in the areas of geography and librarianship:

GEOGRAPHICAL LOCATION CODES
NAMES, GEOGRAPHICAL
GEOGRAPHY--MATHEMATICAL MODELS
PHOTOGRAPHY IN GEOGRAPHY
MEDICAL GEOGRAPHY

LIBRARIES--AUTOMATION
ADVERTISING--LIBRARIES
ACQUISITIONS (LIBRARIES)
BOOKKEEPING, LIBRARIES
LIBRARY ADMINISTRATION
MILITARY LIBRARIES
LIBRARIES, COLLEGE AND UNIVERSITY
LIBRARIES AND AFRO-AMERICANS

Note the different types of headings and inconsistent heading constructs: adjective-noun phrases, prepositional phrases, subdivided headings using hyphens, and inverted headings with commas. Within the same heading form, word order inconsistencies are common. Concepts may appear as both lead term subfield 'a' components and subdivisions. Consider the following examples:

LIBRARIES--FRANCE and PARIS--LIBRARIES
INDIANS OF NORTH AMERICA--ARCHITECTURE and ARCHITECTURE--MEXICO
LAW, ANGLO-SAXON and ROMAN LAW
MORTGAGES (ROMAN LAW) and COMMON LAW--ROMAN INFLUENCES
ABELIAN GROUPS and GROUPS, CONTINUOUS
CHRISTIAN ART AND SYMBOLISM and ART OBJECTS, EARLY CHRISTIAN
ANTIQUITIES--COLLECTION AND PRESERVATION and ALASKA--ANTIQUITIES

In addition, in the 80 years of subject cataloging represented by LCSH, synonymous terms have been established and retained. Consider the following subject headings where the terms "film," "films," "moving pictures," and "motion pictures" are used interchangeably:

WORLD WAR 1939-1945 IN MOTION PICTURES
SUPERNATURAL IN MOVING PICTURES
FANTASY FILM
SCIENCE FICTION FILMS
MOVING PICTURES AND LITERATURE
FILM ADAPTATIONS
RELIGION IN MOTION PICTURES
CLERGY IN MOVING PICTURES
RELIGIOUS FILMS
MOVING PICTURES--CATALOGS
GEOLOGY--FILM CATALOGS

Because of the lack of cross-references and unpredictable phrase constructions, concepts not appearing as lead terms are often difficult to locate. A concept may appear as a subdivision in several different strings, e.g. UNITED STATES--HISTORY--COLONIAL PERIOD--SONGS AND MUSIC; UNITED STATES--SOCIAL LIFE AND CUSTOMS--COLONIAL PERIOD; and BOSTON--HISTORY--COLONIAL PERIOD. A specific topic may appear as a topical subdivision or the object in a prepositional phrase. For example, titles on the Iowa 7th regiment in the civil war are under the subject IOWA--MILITIA--7th REGIMENT. Landscape architects must search through headings such as WATER IN LANDSCAPE ARCHITECTURE, while a specific architectural edifice may be assigned a heading like SAN FRANCISCO--STREETS--GOLDEN GATE FREEWAY. Again, for the subdivided headings there is no cross-reference from the specific topic (e.g., GOLDEN GATE FREEWAY) to the more general heading.

In addition, a topic may appear in certain LCSH as a direct term-pair coordination such as LIBRARIES--AUTOMATION but may also be separated by intervening terms in numerous other relevant LCSH, e.g.,

LIBRARIES--UNITED STATES--AUTOMATION
LIBRARIES--OHIO--AUTOMATION;
LIBRARIES, COLLEGE AND UNIVERSITY--AUTOMATION.

D. Indexing Exhaustivity

Another limitation of LCSH is the lack of indexing exhaustivity, or sparseness of assigned headings. Only 1.4-1.7 subject headings are assigned per work. For instance, a biographical dictionary of Scandinavia is assigned the two headings SCANDINAVIA--BIOGRAPHICAL INFORMATION and FINLAND--BIOGRAPHICAL INFORMATION, but nothing reflects its coverage of Norway, Sweden, Denmark, and Iceland. The title The Insecticide, Herbicide, Fungicide Quick Guide, assigned the single subject heading PESTICIDES--HANDBOOKS, MANUALS, ETC., is another example of this subject headings sparsity.

E. Syndetic Structure

There are numerous instances where reciprocal "see also" references between related headings have not been provided. For example, there is a "see also" from ABDOMEN--DISEASES to ABDOMINAL PAIN but not in the reverse direction. The subject headings ANGLO-SAXONS and GREAT BRITAIN--HISTORY--ANGLO-SAXON PERIOD are not connected by cross-references, nor are the various headings which include the terms MIDDLE AGES or MEDIEVAL, e.g., ART, MEDIEVAL and MIDDLE AGES IN ART.

There are no upward "see also" references from specific to more general terms. As a result, the catalog user is not referred from specific Scandinavian countries to SCANDINAVIA for biographical information. Because an upward

cross-reference from a heading such as ART--ENGLAND to a more general heading like ART--GREAT BRITAIN is not made, users beginning a search at a level too specific will almost certainly miss relevant titles.

In addition, some general cross-references or multiples appear in lieu of a comprehensive set of references. For example, at the term MEDIEVAL there is the instruction SEE ALSO CIVILIZATION, MEDIEVAL; ART, MEDIEVAL; ETC. This forces the medievalist to search the entire alphabet to locate headings such as ART METAL-WORK, MEDIEVAL. Also, a user entering the subject catalog at a subdivided heading such as TRANSPORTATION--CANADA will find no suggested related headings. However, the user entering initially at the non-subdivided heading TRANSPORTATION will be directed to TRANSPORTATION AND STATE; RAILROADS; ETC. And, indeed, the catalog may contain entries such as TRANSPORTATION AND STATE--CANADA and RAILROADS--CANADA. Only at the broadest point in the hierarchy will there be references to narrower or related subject terms.

The problems connected with syndetic structure are made even more acute by the fact that many COM format and online public access catalogs presently do not include the cross-references provided by LCSH.

F. Lack of Guiding Principles in LCSH

Modern controlled-vocabulary direct entry subject schemes--the class to which LCSH belongs--are organized around three characteristic principles. The first feature is that these subject headings systems, whether string languages (PRECIS, chain and faceted schemes) or thesarus-driven (MEDLARS, ERIC) emphasize the assignment of subject terms reflecting all significant subject components, coextensive with the scope of the work. In the case of MEDLARS, 10-12 descriptors are assigned for post-coordinated retrieval. The string languages support multiterm pre-coordinated headings with access provided to all components by a rotation algorithm.[10] The second characteristic feature of these schemes is a provision for a keyword arrangement or index to the indexing vocabulary to facilitate descriptor selection for print users. Of course, the online versions allow component word access. The third common feature of direct-entry subject heading schemes, especially thesauri, is a hierarchical or systematic classified display, with narrower term, broader term, and related term (NT, BT, and RT) relationships identified for optimum retrieval performance.

LC, while espousing the principle of direct entry for subjects, has never completely abandoned in practice a classified approach in its subject catalog. The guiding principles for LCSH were devised by Charles Cutter in a simpler time for a simpler catalog. Writing in 1904, Cutter identified less than 300 subjects consisting of an adjective and noun phrase. Because of this small number, there was no need for rigorous rules concerning citation order in multiword headings.[11] Today, with the greatly expanded number of subject

names, the result is an unsatisfactory hybrid subject heading scheme lacking the consistent arrangement, entry predictability, and syndetic structure of the more pragmatic, controlled-vocabulary direct-entry subject schemes.

III. ENHANCING SUBJECT ACCESS

Discussion of the implementation of a subject access function will be confined to applications for an OCLC-type online architecture using fixed length index records and three-element derived index and search keys--possibly of the type 5, 0-5, 0-2. In the MARC record format, subject headings appear as 6XX fields continuing a single subfield 'a' and other optional subfields.

Two fundamental assumptions will be made. First, solutions to the subject access problem will not include Boolean search techniques. The processing overhead and disk storage requirements connected with the inverted file arrangement to support Boolean search logic would seriously degrade system response in an online environment of some 4,000 terminals. Moreover, past experience with training catalog or data base end-users to wield Boolean operators has shown that many nonexperienced users find Boolean searching complex.[12,13,14] Experienced data base users who are not information specialists tend to conduct fairly simple, straight-forward online bibliographic searches.[15] For these individuals, the types of access points provided by the subject function described in this technical report will be sufficient. It is also assumed that no national system for augmenting or replacing LCSH, as promoted by Pauline Atherton Cochrane, will be instituted and that only data present in the MARC record or directly linked, such as LC classification schedules, will be available as sources of subject information.[16]

A. Rotation of Subject Headings

Clearly, an effective subject retrieval function cannot be based on LCSH alone. Several techniques to improve subject access were investigated in conjunction with two grants awarded to the Iowa State University Library.[17,18] Using data from OCLC MARC archive tapes and a master LC subject authority tape, various phrase rotation algorithms were applied to 6XX fields to generate added index keys.

Providing additional subject access points by manipulation of assigned LCSH can improve subject retrieval in several ways. It can ameliorate the problems connected with inconsistent phrase constructions by providing direct access points to users with the proper terms but non-LCSH word order, and it can bypass some of the weaknesses of the existing syndetic structure by collocating terms such as ANGLO-SAXON; COLONIAL PERIOD and MEDIEVAL which can, in different instances, appear as lead terms, subdividing terms, or subfield 'a' inverted terms. This collocation is particularly important in online environments where browsability is reduced.

This LCSH rearrangement algorithm provides-term pair data element collocation for subject heading subfields and access by significant subject heading words and subfield 'a' term pairs. The rotation procedure is based on a syntactical analysis of LCSH word and subdivision patterns, and is designed to generate the most desirable term combinations, while at the same time, yielding a manageable number of added index keys. The following reorderings comprise the additional headings, from which three-element derived index keys can be extracted:

(a) for LCSH with n>1 subfields, the set of all permutations of the n subfields taken two at a time with the remaining subfields added to the string in sequence order of their position in the original LCSH, and leaving word order within subfields intact;

(b) for LCSH with n>1 subfields, the significant keywords from all non-subfield 'a' subfields, with the remaining subfields added in their original sequence.

(c) for LCSH with n>1 significant subfield 'a' words (processed against a stopword list), the set of all permutations of the n words taken two at a time with the remaining subfields attached in their original sequence to the reordered subfield 'a' elements.

In practice, only selective data in subfields 'a', 'b', 'c', 'y', 'x', 't', and 'z' are reordered and certain high-volume form subdivisions (both $k and $x such as HISTORY, CONGRESSES, DICTIONARIES, etc.) are not allowed lead term status.

To illustrate, the subject heading $a LIBRARIES $z GREAT BRITAIN $x AUTOMATION yields the following headings for index key generation:

1. LIBRARIES--GREAT BRITAIN--AUTOMATION
2. GREAT BRITAIN--AUTOMATION--LIBRARIES
3. AUTOMATION--LIBRARIES--GREAT BRITAIN
4. LIBRARIES--AUTOMATION--GREAT BRITAIN
5. AUTOMATION--GREAT BRITAIN--LIBRARIES
6. GREAT BRITAIN--LIBRARIES--AUTOMATION

Entry 4 collocates the heading keywise with the general subject heading LIBRARIES--AUTOMATION, and with index keys generated from headings such as LIBRARIES--OHIO--AUTOMATION. Entry 3 would be a logical approach to the subject for someone who has examined titles assigned ADVERTISING--LIBRARIES, as would Entry 6 for a user familiar with the subject heading LONDON--LIBRARIES.

As another example, the subject heading $a ART METAL-WORK, MEDIEVAL $z GEORGIA (TRANSCAUCASIA) is rotated to yield the headings:

ART METAL-WORK, MEDIEVAL--GEORGIA (TRANSCAUCASIA)
GEORGIA (TRANSCAUCASIA)--ART METAL-WORK, MEDIEVAL
GEORGIA--ART METAL-WORK, MEDIEVAL
TRANSCAUCASIA--ART METAL-WORK, MEDIEVAL
METAL-WORK--MEDIEVAL--GEORGIA (TRANSCAUCASIA)
ART--MEDIEVAL--GEORGIA (TRANSCAUCASIA)
MEDIEVAL ART--GEORGIA (TRANSCAUCASIA)
MEDIEVAL METAL-WORK--GEORGIA (TRANSCAUCASIA)
METAL-WORK--ART--GEORGIA (TRANSCAUCASIA)

For inverted headings (headings with a comma not followed by a date in subfield 'a'), an additional heading is generated by rotating the inverted terms to the lead position to create a direct entry, and subdividing by any subfields present in the original heading. In the previous example, this rotation produces the heading MEDIEVAL ART METAL-WORK--GEORGIA (TRANSCAUCASIA). For inverted headings with more than one subfield, such as ART METAL-WORK, MEDIEVAL--GEORGIA (TRANSCAUCASIA), an additional customized heading composed of the subfield 'a' terms preceding the comma, subdivided with the remaining subfield(s) is produced: in this case, ART METAL-WORK--GEORGIA (TRANSCAUCASIA). This is to ensure three word collocation between the inverted heading and the reduced LC subject heading ART METAL-WORK--GEORGIA (TRANSCAUCASIA), where no cross-references have been made.

As another example, the heading LIBRARIES, COLLEGE AND UNIVERSITY--AUTOMATION will yield LIBRARIES--AUTOMATION when the inverted form is dropped, so that this heading will be retrieved in response to the search key LIBRA[RIES], AUTOM[ATION].

Several other customized rotation procedures have been designed. For subject headings composed of multiword subfield 'a' data elements with additional subfield(s), the last word in subfield 'a' is combined with the subdivision(s) to produce an added heading. For example, the subject heading ELECTRONIC DIGITAL COMPUTERS--PROGRAMMING is manipulated to yield the added heading COMPUTERS--PROGRAMMING.

With this rotation algorithm, all significant LCSH subfield data (e.g., 7th REGIMENT; COLONIAL PERIOD; GOLDEN GATE FREEWAY) and subfield keywords (e.g., MEDIEVAL; GEOGRAPHY) are brought to the lead position. In addition, generating the pairwise combinations allows truncated retrieval by term-pair coordinations such as LIBRARIES--AUTOMATION, without forcing the user to scan all of the many subject headings containing the keyword LIBRARIES, where the two desired terms often will be separated by intervening words. Because the system is forgiving of word order, these rotation-produced headings allow users to enter one search key and retrieve a display of relevant headings

without having to rekey the string in LCSH word order. This is very important because of the previously mentioned tendency of library patrons to search in only one place in the subject catalog and then to give up, whether they have found the material they are seeking or not.

B. Number of Subject Heading Access Points

Name and subject MARC authority records are organized around an authoritative 1XX heading with associated "see from" 4XX, and "see also from" 5XX fields included. A sample of the subject authority record format is shown in Figure 1. For online searching environments, these fields must be extracted from the authority records and made available as index keys. They are needed as "see" and "see also" cross-references and are displayed in response to search inquiries matching their derived index keys. We selected 9,000 sample records from the 1979 LCSH master authority tape for test processing. For each LCSH subject authority heading there is an average of 1.33 5XX (360) fields, and .7 4XX fields, although many of the 4XX headings conflict with rotated headings and can be suppressed.

The LC subject authority data base contains approximately 90,000 headings. Because of the flexibility built into LCSH by the use of multiples and pattern headings, other LC subject headings (called "nonprint" by LC staff), especially chemical, genus-species, personal, and geographic names will appear in the OLUC. The actual number of LC subject headings may be more than 600,000.

1XX Subject heading, corresponds to 6XX fields in OLUC records.

360 General explanatory "see also" references. Yields headings of the form: "1XX see also 360".

4XX "See from" fields; must be extracted to yield "see" headings of the form: "4XX see 1XX."

5XX "See also from" fields; must be extracted to yield "see also" headings of the form: "5XX see also 1XX."

FIGURE 1. Subject Authority Record Format

A study by O'Neill and Aluri over a sample of the OCLC data base showed that 67% of LCSH in monographic records contained at least one added subfield

(i.e., subfield 'a' and an additional subfield), but only 2.5% contained three added subfields, and only 0.2% contained four or more added subfields.[19] The sample LCSH records were processed according to the above permutation-based algorithm. It was found that for each 1XX heading, five access points or index keys were generated. Assuming (from O'Neill-Aluri) an average 1.32 subject headings per record, the rotation algorithm would generate 6.6 (5 x 1.32) index keys per bibliographic record.

C. Title Word Access

The Iowa State studies investigated the use of title strings to complement subject access. Catalog use and user behavior studies have shown the value of title word access in both known-item and subject retrieval.[20,21,22] An analysis of title, rotated title, and X40 title field strings produced from OCLC archive tapes, using 300 non-fiction monographic records, revealed that some 65% of the records contain at least one significant title string form that is not matched by a subject heading or cross-reference.

The Iowa State studies suggested two possible uses of these title strings. The strings could be used as additional descriptors, that is, added to the dictionary of subject headings. In this way, one could search for imbedded title phrases, and then examine the retrieved bibliographic records to determine the appropriate subject headings to be used in further subject searching. This use of title-word searching to identify useful subject headings has proved to be an effective means of augmenting subject access in the online catalogs of RLIN and the University of California libraries.[23,24]

The other approach, suggested for the subject retrieval function described in this report, employs a title word(s) extraction algorithm to generate additional "see" references to subject headings. For nonfiction, nonjuvenile, nonreference English language type 'a' monographs, 245 (title) field subfields 'a' and 'b' are extracted. These strings are successively cycled or rotated so that all title and subtitle words are given lead term status and then checked against a 600-word stopword list. The surviving title strings are truncated to three significant terms (employing a much shorter stopword list to remove prepositions and other common terms), and a search key is formed and checked for matches against the 1XX/6XX subject authority headings and 4XX references. The unique title keyword strings would be used as additional "see" cross-references to the first 650 topical subject heading assigned to the work. The existence of an assigned 650 topical subject heading, void of a 600 field, is a necessary condition for generating a title-derived "see" reference. This process eliminates a large number of undesirable title strings associated with personal name and geographic subject headings (e.g., FRANZ KAFKA). Generating cross-references from title strings also has the additional advantage of increasing only the size of the subject headings file and index, but not the bibliographic file index.

This title word extraction approach provides a mechanism for retrieving titles on subjects like "steel highway bridges" and "highway bridges" by linking these phrases from the title: Methods of Design for Steel Highway Bridges, with the assigned LC subject heading: BRIDGES, IRON AND STEEL. It is clear that some use of title words will be necessary to retrieve information from search keys based on natural language terms like "space shuttle," "Proposition 13," and "military-industrial complex," since the corresponding LC subject headings: REUSABLE SPACE VEHICLES, REAL PROPERTY TAX--CALIFORNIA; and INDUSTRY AND STATE--UNITED STATES are likely to be less familiar to the typical user. Furthermore, LCSH cross-references may not be provided until years after the term is introduced, or in some cases may never be provided.

For the public access catalogs that do not include the LCSH cross-references, the importance of title word access is paramount. The Library of Congress's slowness in assigning new headings in LCSH is an insidious and continuing problem. As an example, titles on the Cuban missile crisis of 1962 were assigned the heading MILITARY BASES, RUSSIA--CUBA until 1975, when the specific heading CUBAN MISSILE CRISIS, OCT. 1962 was established.

It is estimated that approximately 10-12 title words/phrases will be associated with each subject heading as additional "see" cross-references.

D. Overview of Subject Access Enhancements

The proposed system for expanding subject access is centered around the creation of additional subject index keys by LCSH phrase manipulation and title phrase extraction. However, generating additional subject keys through rotation techniques will not solve the problems in LCSH that are vocabulary related (e.g., "film," "motion picture," "moving picture") or caused by the sparseness of assigned headings, the lack of adequate cross-references, the presence of archaic heading forms, and the assignment of general headings with high postings. Some of these situations can be ameliorated through the use of title word strings to augment the syndetic structure of cross-references. With LC now in a position to exercise greater vocabulary authority control, the future may bring an improvement both in the breadth of LC's assignment of subject headings, and in the quality of its subject analysis.

The rotation procedures combined with title keyword/keyphrase access provide a subject retrieval function that, in some measure, simulates a free-text Boolean search capability over 6XX and title fields. This type of access-point file would be somewhat new to OCLC, but is characteristic of systems maintained by commercial online data base vendors (Lockheed, BRS, SDC). These vendors recognize the complementary nature of controlled vocabulary access and natural language free-text searching. The expanded subject access points described here provide greatly increased subject retrieval over the traditional card catalog.

E. Other Measures to Improve Subject Access

Other refinements to LCSH could be examined. In particular, there are five software-driven enhancements that would place LCSH more in line with the principles that characterize modern subject analysis.

(1) The generation of reciprocal "see also" references between headings not characterized by BT or NT relationships. This could be accomplished by comparing a heading's "see also" references with its 5XX fields.

(2) The reformatting of the subject scheme into a thesaurus-style display, similar to ERIC's, with keyword index access.

(3) The use of the BT designation byte in the control w subfield of 5XX authority fields to produce upward NT to BT "see also" references from authority headings. This would produce a reference from ART--ENGLAND to ART--GREAT BRITAIN. However, this cross-reference structure is only available in a machine-readable format for headings in the LCSH subject authority file. Subject headings in the OLUC formed by the application of free-floating or geographic subdivisions have no accompanying cross-reference fields. For example, the heading ART--WALES, which is not in the LCSH subject authority file, will have no associated cross-references to or from ART--GREAT BRITAIN or ART--ENGLAND. To remedy these shortcomings, OCLC might allow member libraries to enhance subject authority headings obtained from the OLUC with the appropriate cross-reference terms.

(4) The use of LC classification schedule terms linked to LCSH by the 053 field in subject authority records.

(5) The linking of LC subject headings to descriptors taken from subject-specific thesauri, using a switching language mechanism.

IV. SYSTEM DESIGN FEATURES

A derived search key-based subject retrieval function, offering improved subject access over traditional card catalog approaches, and organized around the previously mentioned LCSH rotation and title word access procedures, with familiar OCLC software conventions, will be outlined in this section. This subject retrieval function would be designed around a 5, 0-5, 0-2 index/search key and would feature subject authority control. This function would allow limiting search output by holding library, publication date, bibliographic type, and language. While several file designs are possible, from the user's standpoint, the system would appear as a subject headings file, a cross-references file, and a bibliographic records file. All three files would be linked and thus would be accessible from each of the others.

A. File Design

While several index file arrangements are possible, the optimum approach would be to partition the files for a two-step subject retrieval function. This two-step process would require a search of an intermediate subject headings index file preceding the retrieval of the bibliographic records. This two-step approach allows examination of candidate subject headings for user selection and search refinement.

The initial search would scan an index to the 600,000 subject authority records. As outlined in Figure 2, this index would contain derived search keys for the subject headings, rotated subject headings, 4XX references, and title phrases associated with the headings. For each subject authority record, there are index keys for the heading and an average of five rotated headings, .7 4XX references, and ten title phrases totaling approximately 16 access points. The index keys also contain an LCSH number pointing to the subject authority records file arranged in LCSH sequence. After retrieval of the candidate subject authority records and selection of a subject heading, the bibliographic records would be retrieved from an index having LCSH number keys with associated OCLC document numbers and qualifying matrix. A separate cross-reference file containing the "see also" references associated with a heading and arranged in LCSH number sequence is also provided. In effect, as diagrammed in Figure 3, the file arrangement provides indirect added access points by increasing the number of stored keys for each subject heading assigned to a bibliographic record. This type of file arrangement raises the number of subject access points assigned to a work from 1.4 to over 22.

Subject authority records file.
Variable length records, average 400 bytes in length. 600,000 records taken from subject authority tape and 6XX fields in OLUC.

Keys for subject authority file. 6,000,000 fixed length records generated from headings, rotated headings, 4XX tracings in subject authority records, and title phrases associated with headings. Approximately 16 access points or index keys for each subject heading.

K(H1)	AN1	F4
K(R1H1)	AN1	F4
K(R2H1)	AN1	F4
K(R3H1)	AN1	F4
K(ST1H1)	AN1	F4
K(TP1H1)	AN1	F4
⋮		
K(TPnH1)	AN1	F4

Authority records contain 1XX headings from authority tape and 6XX fields in OLUC, 4XX and 5XX tracings fields and 9XX title phrase fields.

I1	AN1	A1	O1
	ANj	Aj	Oj

Bibliographic file index fixed length. 9.8 million records.

AN1	B1	M1
AN1	B2	M2
⋮		
AN1	Bj-1	Mj-1
AN1	Bj	Mj

Cross-reference index, keyed on extracted 5XX tracings, connected to 1XX headings from subject authority tape

ANj	AN1
ANj	ANk
ANj	AN1
⋮	
ANk	ANm
ANk	ANj

A1 is subject authority record;
I1 is flag for existence of authority record heading (H1) in OLUC;
AN1 is pointer/identifier for subject authority record;
K(R1H1) is key generated from 1th rotation of heading H1 (1XX field in authority record A1)
F4 is 4XX flag for authority file keys to indicate key derived from "see to" reference or title phrase
B1 is OCLC# of bib record REC1
O1 is number of postings in bib database for heading H1
M1 is limiting matrix (language, year of publication, bib type, etc.) for bib record REC1
ST1H1 is 1th 4XX tracing in authority record A1
TP1H1 is title phrase TP1 associated with heading H1

FIGURE 2. Subject Authority File

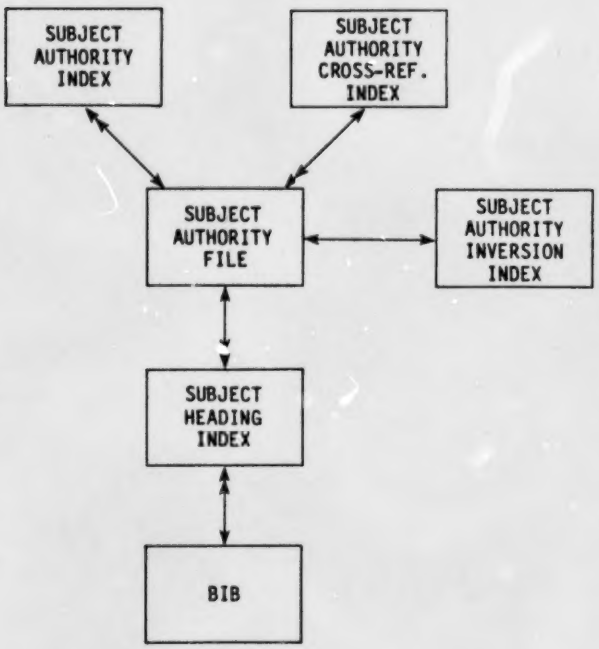

FIGURE 3. Subject Access Data Base

B. Subject Search Transactions

A typical subject search transaction would involve a multi-step process beginning with the user searching a subject headings file or dictionary, made up of 6XX fields from the OLUC, and 1XX headings and 4XX fields from the LC subject authority master tape(s). From the truncated/collective display of subject headings and 4XX "see" references, the user would select a desired heading or reenter another subject heading search key. Linked to each subject heading is the bibliographic file of OLUC records assigned to the selected heading, and a cross-reference file of other associated subject headings linked by a "see also" reference. After selecting a desired subject heading a user can display either the associated bibliographic records or cross-references. In turn, a heading selected from the cross-reference file display would retrieve an associated display of bibliographic records assigned

to the subject heading. A command language structure would allow switching between the subject heading displays and the associated bibliographic record displays. Figure 4 depicts a user transaction in the form of a flow diagram for the subject retrieval function.

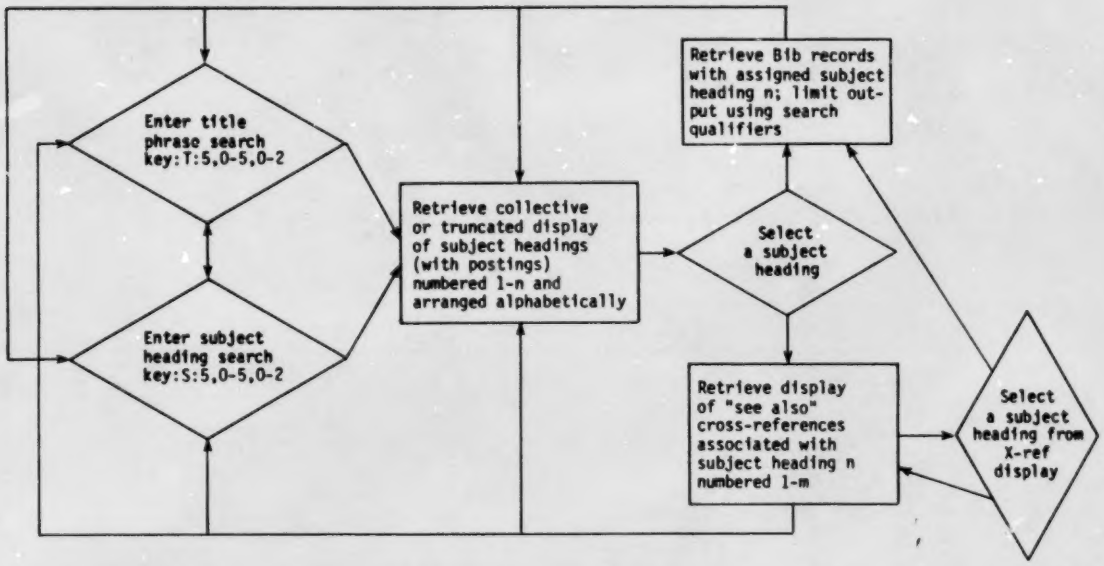
Enter title phrase search key:T:5,0-5,0-2
Enter subject heading search key:S:5,0-5,0-2
Retrieve collective or truncated display of subject headings (with postings) numbered 1-n and arranged alphabetically
Select a subject heading
Retrieve Bib records with assigned subject heading n; limit output using search qualifiers
Retrieve display of "see also" cross-references associated with subject heading n numbered 1-m
Select a subject heading from X-ref display

FIGURE 4. Subject Retrieval Function

As an illustrative example, the search inquiry "S:COLON, PERIO," would yield a numbered display of subject headings and 4XX references matching the search key, such as:

SUBJECT HEADINGS AND SEE REFERENCES

1. AMERICAN DRAMA--COLONIAL PERIOD
2. AMERICAN POETRY--COLONIAL PERIOD
3. BOSTON--HISTORY--COLONIAL PERIOD
4. BOSTON--POLITICS AND GOVERNMENT--COLONIAL PERIOD
5. FURNITURE, COLONIAL (USED FOR: COLONIAL PERIOD FURNITURE)
6. UNITED STATES--HISTORY--COLONIAL PERIOD
7. UNITED STATES--HISTORY--COLONIAL PERIOD--SONGS AND MUSIC
8. UNITED STATES--FOREIGN RELATIONS--COLONIAL PERIOD
9. UNITED STATES--SOCIAL LIFE AND CUSTOMS--COLONIAL PERIOD

From this display, the user can reenter a search key or select a line number (subject heading) and ask for the associated bibliographic records, "see also" cross-references display, or subject authority record. The request for cross-references from heading 6 would produce a display such as:

CROSS-REFERENCES FROM SELECTED SUBJECT HEADING:

1. UNITED STATES--HISTORY--COLONIAL PERIOD

SEE ALSO

2. GREAT BRITAIN--COLONIES--AMERICAN
3. PEQUOT WAR, 1636-1638
4. STAMP ACT, 1765

From here, the user can select the bib records associated with a selected cross-reference, go back to retrieve bib records from the originally selected heading, receive the display of cross-references associated with a selected heading, or review the previous subject headings display. As outlined here, the user could "jump" across cross-reference displays but would not be able to jump back to previous cross-reference displays.

Displaying occurrences or postings in conjunction with the subject headings and cross-reference headings would be a desirable modification of this scheme. That is, the subject heading display would include the number of bibliographic record postings associated with each heading to assist users in selecting appropriate headings.

C. Limiting of Subject Search Results

The bibliographic record indexes would be constructed using packed bit matrices and one-byte codes. This design would support the limiting of bibliographic record search output within the indexes themselves, without the need for accessing the full records.

Because users typically will not be conducting known-item searches, multiple citations will usually be required to satisfy a search need. A screen of truncated/collective bibliographic entries may well contain all relevant titles. However, the user will not want to examine--and OCLC will be unable to store in work buffers and display--thousands of records assigned to a selected subject heading. Yet, it will be necessary to ensure that, at some point, less than 1500 hits are retrieved.

To ensure this limited number hits, searches could be qualified by year and language of publication; bibliographic level and type; the presence of illustrations, maps, etc.; or by the publication form (e.g., government publication, thesis or dissertation, microform, and other data from the 008 fixed field). These additional qualifiers could be provided for all types of OLUC searches. Perhaps some initial holdings information (e.g., states, regions, whether LC MARC record) could also be stored in the index records. It may also be desirable to allow further limiting of bibliographic search results to holding library or specified library group by screening the institution symbol display string. The typical subject retrieval function user will want the most recently printed monographic titles in English (no reprints, microforms, juvenile literature, etc.) and only those titles available in his/her library system or network. These default user parameters could be codified and placed in a pre-stored profile, linked to the user's logon number. This profile could be overridden when needed, after viewing results at the search stage.

Even for searches qualified by holding library, date, language, bibliographic type, etc., there may be cases where more than 1500 items are retrieved. There also will be many instances of highly posted items, i.e., items with hundreds of references.

To simplify the search and display process, it might be desirable to set an online display limit of 50-100 items or 2-4 collective screens for fully qualified search results. In this way, the typical search transaction would result in the user's examining the most recent 50 to 100 titles, in reverse chronological order. These are the titles that are most likely to be relevant for a user's search needs. In effect, by using the imprint dates, this limit operation would further qualify search results to some arbitrary number of retrieved titles.

Some vastly encompassing subject headings exist (e.g., EDUCATION, with 30,000 postings). Searching through entire sets like these for even the 100 most-recent titles by a holding library would commit a large amount of processing time. However, a more efficient method might be to secondarily sort bibliographic index records by date, following the practice of the commercial online data base vendors. This procedure would expedite system processing and the throughput of search inquiries and responses.

We might also consider displaying abbreviated format online records, allowing the truncated display format to include OCLC number, 050, 1XX, 245, 260, 5XX, and 6XX fields. This display would allow all necessary information for title retrieval of multiple records to be shown on one screen.

D. Charging for Offline Products and Basic Service

Providing offline batch output on a transaction fee basis would be an additional service for users who desire merged customized bibliographies or formatted paper output of subject search results. There is ample precedent among libraries for paying for offline printouts of subject bibliographies. This system could provide title-level availability because of its capability to limit subject search results to holding libraries.

The subject retrieval component could be charged for on a license basis, with licensed institutions paying a yearly fee for unlimited searching. This fee could be based on a sliding scale tied to an institution's annual first time use and first time update of a record (FTU's/FUP's).

E. Public Access Enhancements

The search inquiry phase of the retrieval process could be made more user-cordial by allowing natural language search inquiries for parsing by software, removal of prepositions, generation of derived search keys, analysis of syntactical structure, etc. Here I would expect a more dialogue-driven approach, allowing multiple line selection of descriptors and more detailed message and instructional displays. The users could be asked to enter a statement of "aboutness", that is, they would be required to initiate the search transaction by completing the phrase "I am interested in works that are about _____". The system would then process the request, formulate search keys, and present a candidate list of subject headings for user examination. A user-entered subject "aboutness" string, such as "the Byzantine influences on Medieval art" would be first stripped of articles, prepositions, and nonsignificant words. From the remaining terms, various term combinations--both term triples and term pairs--would be system-generated and searched against the recorded subject heading index keys. The resulting list of retrieved subject headings (perhaps ranked by degree of match) then would be retrieved and displayed for user selection of the desired subject heading(s). Response to the "aboutness" string in the aforementioned example would include the retrieved subject headings:

1. ART, MEDIEVAL--BYZANTINE INFLUENCES
2. ART, EUROPEAN--BYZANTINE INFLUENCES
3. ART, BYZANTINE
4. ART, MEDIEVAL

The bibliographic records and/or "see also" cross-references associated with the selected subject heading(s) then would be retrieved and displayed in the same manner as outlined earlier.

F. Manipulation of Search Keys

User-entered search keys which fail to retrieve any candidate headings would be system manipulated to yield reduced search keys and then checked against the subject headings index file again. That is, without additional keying, search key arguments for failed searches would be selectively reordered and truncated to yield alternative search keys. The user could be informed that the initial search key failed to retrieve any candidate headings and that further searching with alternative keys (displayed for the user's information) was taking place.

For example, a user interested in titles on the history of the southwestern UNITED STATES might enter a search key of "UNITED STATES, SOUTH, HI" to retrieve a list of candidate subject headings. However, the appropriate LC headings for works on the southwestern UNITED STATES are SOUTHWEST, NEW and SOUTHWEST, OLD. No user-entered search statement (whether Boolean, string match, or derived key) which contains the term "UNITED STATES" will retrieve the relevant subject headings because the term "UNITED STATES" does not appear in the heading. The iterative search statement truncation algorithm specified in this proposed subject retrieval function would produce the reduced key "SOUTH,____, HI". From this, a candidate list of subject headings would include: SOUTHWEST, NEW--HISTORY and SOUTHWEST, OLD--HISTORY.

G. Comparison with Boolean Searching

A direct comparison between the subject approach suggested in this report and Boolean search systems is difficult, primarily because the access points in Boolean-based systems are derived a posteriori as a function of the command operators.

The subject approach described in this report simulates a Boolean search capability over MARC record 6XX fields by providing selected precoordinated term permutations of data elements from subject heading fields. In this way, Boolean term permutations are built into the system indexes. Subject heading keywords provide additional access points.

Moreover, the approach is based on a syntactical analysis of LCSH and title word phrases which identify the most desirable term coordinations for subject retrieval. While the proposed function does not possess the full term coordination capabilities of a Boolean search system, it does provide some retrieval advantages over Boolean search systems. Boolean-based search systems are plagued with retrieval problems connected with false word coordinations--both within single subject headings and because of term crossover between different headings or different fields (e.g., title and subject headings). For example, a search for "English and History" with the intention of retrieving titles on English history will retrieve works assigned the subject heading ENGLISH POETRY--EARLY MODERN, 1500-1700--HISTORY AND CRITICISM. A search for the topic "English and literature" will retrieve titles assigned the pair of subject headings ENGLISH POETRY and SHEPHERDS IN LITERATURE. The approach detailed in this technical report avoids these types of problems by not coordinating the terms English and history, or English and literature in these precoordinated subject heading index keys.

The proposed non-Boolean retrieval function supplies system users with a candidate list of subject headings. Therefore the subject search is first performed over an intermediate index of subject headings. This allows users to more easily refine the search and focus on the specific heading(s) relevant to their search needs. The operational Boolean systems typically return document citations in response to search arguments or require users to employ a special command to invoke an authority file. This can be a problem both when zero postings are indicated and when, because of false term coordination, a large number of hits are retrieved for display. Separating title words from subject heading terms also helps prohibit the occurrence of false term combinations.

Clearly, searches based on complicated logical expressions can be more readily accomplished within a Boolean framework. But, how often these complex searches will be required or utilized is an open question. Also, effectively constructing a complex search argument, based on the meager subject data elements present in the MARC record, offers the searcher another problem.

One possible option in the development of a subject retrieval function for the OLUC would be to consider the non-Boolean selective term permutation approach described in this report as an intermediate phase prior to the implementation of a full Boolean capability, utilized in a distributed network environment.

H. Summary of Subject Retrieval Function

The approach described in this report would provide a subject retrieval function centered around the creation of additional subject access points and cross-references from assigned LCSH and title strings. The system would employ familiar OCLC searching-conventions and commands, and is designed to simulate a Boolean search capability by building precoordinated term

combinations into the derived index keys and providing access through subject heading keywords and subfield 'a' term pairs. This flexibility alleviates many of the problems of LCSH word order construction, which are unresolvable by the application of phrase order rules and impossible for librarians and other users to predict. The two-step file design outlined earlier provides expanded access to the individual subject headings assigned to records in the OLUC. This approach allows an increase in the average number of subject access points per title from 1.4 to over 22, but not at the expense of an increased bibliographic records index file. In addition, the file design supports a mechanism for subject authority control, accommodating subject heading changes that LC will institute. The file design would preclude having to immediately update the entire OLUC to reflect the changes.

This subject retrieval function requires users to examine an intermediate index of candidate subject headings, allowing greater search refinement than is possible in Boolean search systems that retrieve bibliographic records. It also features a provision for system manipulation of search arguments which fail to retrieve any candidate headings by reducing terms for subsequent re-searching of the subject headings index.

The subject retrieval function would be used primarily by library staff, especially public service librarians and original catalogers. The typical user search will seek the most recent titles assigned a specific subject heading. For this reason, and because of the high number of postings associated with many subject headings, a secondary key arrangement by imprint date for the bibliographic file index would optimize system processing of requests.

The system will also be useful as a general reference tool (e.g., to obtain a list of Roman emperors from the 600 tag, subfield 'c', EMPEROR OF ROME rotation). The 5, 0-5, 0-2 derived key for searching the subject headings and 4XX field dictionary would allow keyword access. Specifying a 5XX field cross-reference file would provide online users with the syndetic structure that is built into LCSH, but not provided by most online and COM format subject retrieval systems.

V. CONCLUSION

The treatment discussed in this report would provide OCLC users with an expanded subject retrieval function on a scale which is unparalleled by commercial data base vendors, which allow a maximum of approximately 100 users connected to a system at one time. This design also provides a retrieval capability vastly improved over the online public access catalogs which permit only exact-match string searching of LCSH, (e.g., those of Northwestern University, CLSI touch-terminal catalogs, Ohio State's LCS system, and the ULISYS system being used at Mission College and other libraries). It offers much greater flexibility in search term word order than the RLIN II subject phrase index. In addition, implementation of this type of system would make OCLC the only bibliographic network offering the online syndetic structure of LCSH (with enhanced title-derived "see" references) and the capacity for subject authority control.

VI. REFERENCES

[1] Minker, Jack and Rosenfeld, Samuel; Introduction and perspectives for the 1971 ACM information storage and retrieval symposium. ACM Symposium on Information Storage and Retrieval, College Park, Maryland, 1-2 April 1971. Proceedings. 1-3. 1971.

[2] Markey, Karen. Analytical review of catalog use studies. OCLC, Inc., Columbus OH: OCLC/OPR/RR-80/2 1980 February; 22.

[3] Dwyer, James R.; The effect of closed catalogs on public access. Library Resources and Technical Services. 25:189; 1981 April/June.

[4] Chan, Lois Mai; Library of Congress subject headings. Littleton, CO: Libraries Unlimited, Inc.; 17-20; 1978.

[5] Gorman, Michael; How the machine may yet save LCSH. American Libraries. 11:557-558; 1980 October.

[6] Freedman, Maurice J.; Malinconico, Michael S., editors; The nature and future of the catalog. Information Science and Automation Division's 1975 and 1977 Institutes on the catalog proceedings. Phoenix, AZ: Oryx Press; 20-33, 89-101, 120-121, 225-256; 1979.

[7] Aveney, Brian; Facing our technology. Journal of Library Automation. 8:85-86; 1975 June.

[8] Atherton, Pauline; Books are for Use: Final report of the Subject Access Project to the Council on Library Resources. Syracuse, NY: Syracuse University School of Information Studies; 1-2; 1978.

[9] Library of Congress. Cataloging Service Bulletin; no. 1. 19; 1978 Summer.

[10] Keen, E. Michael; On the generation and searching of entries in printed subject indexes. Journal of Documentation. 33:15-45; 1977 March.

[11] Cutter, Charles A.; Rules for a Dictionary Catalogue. Government Printing Office. Washington D.C.; 72; 1904.

[12] Kugel, Peter; Dirty Boole? Journal of the American Society for Information Science. 22:293-294; 1971 July-August.

[13] Summit, Roger K.; Popular illusions relating to the costs of online services. Paper read at the 1980 American Library Association Annual Meeting, June 28-July 4, 1980. Quoted in American Libraries, 425; 1980 July-August.

[14] Summit, Roger K.; Dialog and the user--an evaluation of the user interface with a major online retrieval system. Interactive Bibliographic Search: The User/Computer Interface. Edited by Donald E. Walker. Montvale, NJ; AFIPS Press; 92-93; 1971.

[15] Fenichel, Carol.; The process of searching online bibliographic data bases; a review of research. Library Research. 2:107-127; 1980-81.

[16] Atherton, Pauline; Revise subject access, too. American Libraries. 12:122; 1981 March.

[17] Mischo, William H.; Expanded subject access to reference collection materials. Journal of Library Automation. 12:338-354; 1979 December.

[18] Mischo, William H.; Expanded subject access to library collections using computer-assisted indexing techniques. 43rd ASIS Annual Meeting. Proceedings 17:155-157; 1980 October.

[19] O'Neill, Edward; Aluri, Rao. Subject heading patterns in OCLC monographic records. OCLC, Inc., Columbus OH: OCLC/OPR/RR-80/2; 1979 August.

[20] Atherton, Pauline; Catalog users' access from the researcher's viewpoint: past and present research which could affect library catalog design. Closing the Catalog. Phoenix, AZ: Oryx Press; 107-115; 1980.

[21] Weintraub, D. Kathryn; The essentials or desiderata of the bibliographic record as discovered by research. Library Resources and Technical Services. 23:291-405; 1979 Fall.

[22] Kaske, Neal K.; Sanders, Nancy P.; Online subject access: The human side of the problem. RQ. 20:56-58; 1980 Fall.

[23] International symposium on the union catalog--a summary report. RTSD Newsletter. 6:38; 1981 July-August.

[24] Miller, Julia; et al.; RLIN at the reference desk. RLIN News. 3/4:13-15; 1981 March-June.